Supporting Information

Schug et al. 10.1073/pnas.0912100106

SI Text

Biological Background. ***Two-component systems and phosphorelays.*** In two-component signaling (TCS), a sensor histidine kinase (SK) and a response regulator/transcription factor protein (RR) connect an input signal to an appropriate response via a transphosphorylation reaction between the two proteins (1). In this reaction a phosphoryl group is transferred from a histidine on the SK to an aspartate residue on the RR. This results in activation of the transcription factor activity of the latter. A signal serves to modulate the autokinase activity of the SK and thereby controls the flux of phosphoryl groups through the system (2). Autokinase activity of the SK is mediated by two domains, a dimeric four-helix bundle domain termed HisKA, which contains the phosphorylatable histidine residue as well as an ATP-binding ATPase domain.

A more extensive version of this signal transduction pathway is the phosphorelay. In these systems, phosphoryl group flux between the input SK and the terminal RR is mediated via additional proteins, whose sole apparent purpose it is to hand the phosphoryl group from one to the next protein following a His-Asp-His-Asp cascade (1). A common misperception (recent, e.g., ref. 3) in the two-component system field is that all phosphorelays are structurally identical. Instead, two structurally distinct versions of the phosphorelay have been described (Fig. S1).

The more common one is not only found in bacteria but also in some eukaryotes and involves a SK, a central single domain RR, an Hpt phosphotransferase domain and finally the RR/transcription factor protein (4). The Hpt domain is structurally distinct from the HisKA domain, forming a monomeric four-helix bundle. A similar domain is also used in the chemotaxis-signaling pathway, which features the atypical five-domain histidine kinase CheA (5).

A second version of the phosphorelay also utilizes the SK and two RR proteins. Instead of an Hpt domain however, the second his-containing protein is a structural homologue of the SK but having lost the ability to bind ATP and autophosphorylate (6, 7). In principal, one can think of this type of phosphorelay as two TCS positioned in tandem, with the second kinase having lost its autophosphorylation ability and having adopted a phosphotransferase role. Such a system is exemplified by the sporulation phosphorelay (8). Here, five individual kinases termed KinA-KinE are known to phosphorylate the single domain RR Spo0F, which passes the phosphoryl group to Spo0B and finally to Spo0A, the terminal transcription factor/response regulator, which initiates sporulation (9, 10).

Structural Similarities and Differences between Spo0B and Sensor Histidine Kinases. The Spo0B phosphotransfer protein in the sporulation pathway shares a likely common evolutionary origin with the typical SK (11). As such it features a central four-helix bundle (the HisKA equivalent domain), which contains the phosphorylatable histidine (12). As in the SK, the four-helix bundle is formed by two individual polypetide chains, each monomer contributing two helices. A rudimentary ATPase domain underscores the common evolutionary origin with true SK proteins, despite Spo0B having lost the ability to bind ATP and autophosphorylate the histidine on the four-helix bundle (7, 11) (Fig. S1). The similarities have contributed to the consideration of Spo0B in complex with the RR Spo0F as an adequate and to date only model structure for the SK/RR complex (11). This structure has since been used extensively as a reference for interpretation of experimental as well as computational results, aimed at identifying surface residues as well as residues involved in interaction specificity of two-component systems proteins (13–17).

Structures of the typical TM0853 SK from *Thermotoga maritima* (the best available X-ray diffraction SK structure) and Spo0B have been superficially compared (18). A detailed structural comparison to support the notion of Spo0B/Spo0F as good albeit not ideal structural model for SK/RR interaction will follow below.

The dimeric four-helix bundle of SK proteins and Spo0B feature two pairs of helical interfaces. These are an intramolecular interface pair formed between the two helices on each monomer (i.e., $\alpha1$ with $\alpha2$ and $\alpha1'$ with $\alpha2'$) and an intermolecular interface pair formed between one helix from each monomer within the dimer (i.e., $\alpha1$ with $\alpha2'$ and $\alpha1'$ with $\alpha2$). When overlaying the catalytic histidine residue along with surrounding residues on helix $\alpha1$, structural similarities and differences between the Spo0B four-helix bundle and the TM0853 four helix-bundle become immediately apparent (Fig. S2*A*). Whereas the two proteins retain the same intermolecular orientation between $\alpha1$ and $\alpha2$‘ (and $\alpha2$ with $\alpha1'$), the intramolecular orientation between $\alpha1$ and $\alpha2$ (and $\alpha1'$ with $\alpha2$‘) differs substantially (Fig. S2*B*). In Spo0B, helix $\alpha2$ does not contact the bound Spo0F RR in the existing cocrystal structure and most primary contacts are made through helix $\alpha1$. For this reason, structural alignment of $\alpha2$ of Spo0B with $\alpha2$ of TM0853 is not possible. Our direct coupling analysis (DCA) performed on SK/RR interaction partners identified 5 contact positions in SK helix $\alpha1$ and 3 in helix $\alpha2$, all connecting to five residue positions in the RR $\alpha1$-helix (15). These contacts were the input for our docking simulations described in the main text. The identified $\alpha2$ contacts had to be ignored for the above structural reason when assembling the Spo0B/Spo0F structure.

To form an active site as a requirement for phosphoryl group transfer, the orientation between $\alpha1$ of a true SK with $\alpha1$ of the RR has to remain conserved with respect to what can be observed in the Spo0B/Spo0F cocrystal structure. The different orientation of helix $\alpha2$ in the SK results in possible additional contacts with RR helix $\alpha1$, consistent with our direct coupling analysis and existing experimental results (15, 17, 19). For this reason, contacts in both helices were used for docking simulations of SK TM0853 with its RR TM0468.

Perhaps more interesting than the described dissimilarity is the structural similarity of the intermolecular surface formed between $\alpha1$ and $\alpha2'$ and $\alpha1'$ with $\alpha2$. In the Spo0B/Spo0F cocrystal structure contacts are made with this interface that were implied in sealing of the phosphotransfer active site from solvent access (6, 7, 11). In particular contacts between Spo0B residues K63 and K67 in helix $\alpha2$ and Spo0F residue Y84 in the ($\beta4$-$\alpha4$)-loop have been experimentally validated to be important (19) (Fig. S2*C*). The observed structural conservation of this interface between Spo0B and SK TM0853 implies that true SK proteins interact with their paired RR in a similar manner. If this were true, one would have expected that our DCA would find some highly correlated residues at this interface, which is not the case. DCA does not pick up conserved residues, but the residues of interest are variable. Another possibility is that the contacts in this region of the protein can be made in a number of different ways, i.e., they differ from SK/RR pair to another SK/RR pair. This would result in dilution of the correlation signal. Consistent with this

notion, the (β4-α4)-loop/helix α4 region of RR, involved in these contacts, is known to be highly dynamic (4, 13, 20).

In summary, structural analysis of Spo0B in comparison with SK TM0853 revealed that the Spo0B/Spo0F structure is indeed an adequate model structure for SK/RR interaction. The important differences are found in contacts made between helix α2 with the RR, which are not realized in Spo0B and could hence were not forthcoming from this structure.

1. Hoch JA (2000) Two-component and phosphorelay signal transduction. *Curr Opin Microbiol* 3:165–170.
2. Szurmant H, White RA, Hoch JA (2007) Sensor complexes regulating two-component signal transduction. *Curr Opin Struct Biol* 17:706–715.
3. Jagadeesen S, Mann P, Schink CW, Higgs PI (2009) A novel ''four-component'' two component signal transduction mechanism regulates developmental progression in *Myxococcus xanthus*. *J Biol Chem* 284:21435–21445.
4. West AH, Stock AM (2001) Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem Sci* 26:369–376.
5. Szurmant H, Ordal GW (2004) Diversity in chemotaxis mechanisms among the bacteria and archaea. *Microbiol Mol Biol Rev* 68:301–319.
6. Varughese KI, Tsigelny I, Zhao H (2006) The crystal structure of beryllofluoride Spo0F in complex with the phosphotransferase Spo0B represents a phosphotransfer pretransition state. *J Bacteriol* 188:4970–4977.
7. Zapf J, Sen U, Madhusudan, Hoch JA, Varughese KI (2000) A transient interaction between two phosphorelay proteins trapped in a crystal lattice reveals the mechanism of molecular recognition and phosphotransfer in signal transduction. *Structure (London)* 8:851–862.
8. Fabret C, Feher VA, Hoch JA (1999) Two-component signal transduction in *Bacillus subtilis*: How one organism sees its world. *J Bacteriol* 181:1975–1983.
9. Burbulys D, Trach KA, Hoch JA (1991) Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay. *Cell* 64:545–552.
10. Jiang M, Shao W, Perego M, Hoch JA (2000) Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*. *Mol Microbiol* 38:535–542.
11. Hoch JA, Varughese KI (2001) Keeping signals straight in phosphorelay signal transduction. *J Bacteriol* 183:4941–4949.
12. Varughese KI, Madhusudan, Zhou XZ, Whiteley JM, Hoch JA (1998) Formation of a novel four-helix bundle and molecular recognition sites by dimerization of a response regulator phosphotransferase. *Mol Cell* 2:485–493.
13. Szurmant H, et al. (2008) Co-evolving motions at protein–protein interfaces of two-component signaling systems identified by covariance analysis. *Biochemistry* 47:7782–7784.
14. White RA, Szurmant H, Hoch JA, Hwa T (2007) Features of protein–protein interactions in two-component signaling deduced from genomic libraries. *Methods Enzymol* 422:75–101.
15. Weigt M, White RA, Szurmant H, Hoch JA, Hwa T (2009) Identification of direct residue contacts in protein–protein interaction by message passing. *Proc Natl Acad Sci USA* 106:67–72.
16. Burger L, van Nimwegen E (2008) Accurate prediction of protein–protein interactions from sequence alignments using a Bayesian method. *Mol Syst Biol* 4:165.
17. Skerker JM, et al. (2008) Rewiring the specificity of two-component signal transduction systems. *Cell* 133:1043–1054.
18. Marina A, Waldburger CD, Hendrickson WA (2005) Structure of the entire cytoplasmic portion of a sensor histidine-kinase protein. *EMBO J* 24:4247–4259.
19. Tzeng YL, Hoch JA (1997) Molecular recognition in signal transduction: The interaction surfaces of the Spo0F response regulator with its cognate phosphorelay proteins revealed by alanine scanning mutagenesis. *J Mol Biol* 272:200–212.
20. Feher VA, et al. (1997) High-resolution NMR structure and backbone dynamics of the *Bacillus subtilis* response regulator, Spo0F: Implications for phosphorylation and molecular recognition. *Biochemistry* 36:10015–10025.

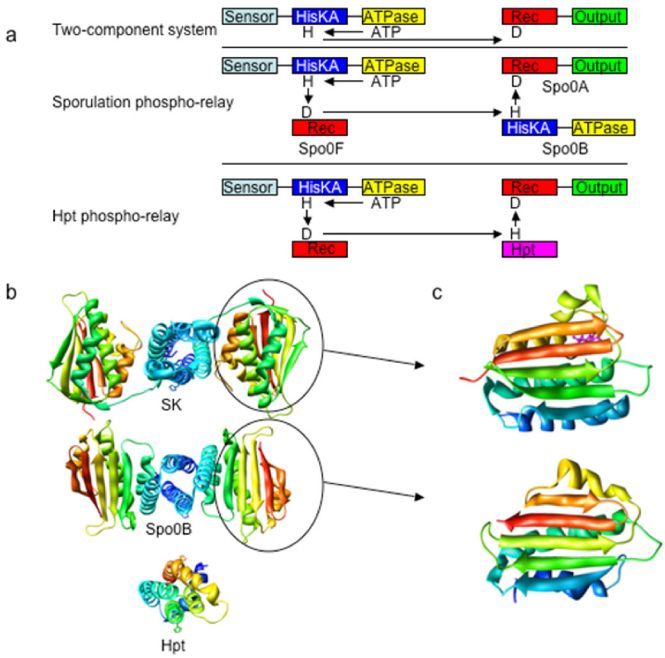


Fig. S1. Two-component systems and phosphorelays. (*A*) Two-component signal transduction systems feature a sensor kinase and a response regulator protein. A signal is sensed by sensor domain (light blue) resulting in phosphorylation of the HisKA four-helix bundle domain (blue) via the catalytic ATPase domain (yellow). An output is created by the effector domain, usually a DNA-binding domain (green), which is regulated by the receiver domain Rec (red). The message between the two proteins is passed by transfer of a phosphoryl group between a histidine on the HisKA domain and an aspartate residue on the Rec domain. An extended version of the two-component system is the so-called phosphorelay, using two additional phosphotransfer proteins. Two structurally distinct version of the phosphorelay have been described. The sporulation phosphorelay features a central Rec protein Spo0F and a SK analogous protein Spo0B, which does not possess the ability to autophosphorylate. The Hpt phosphorelay features a structurally distinct Hpt domain protein, instead of a second SK-like protein. (*B*) Structural comparison of the SK TM0853 with Spo0B and an Hpt protein. Whereas the Hpt domain is a monomeric single domain protein, both SK and Spo0B share a two-domain dimeric architecture, demonstrating Spo0B's evolutionary relationship to SK proteins, rather than to Hpt domains. (*C*) Comparison of the C-terminal ATPase domain of SK and Spo0B demonstrates the similarity in structure, despite the Spo0B domain having lost the ability to bind ATP and autophosphorylate the Spo0B four-helix bundle histidine residue.

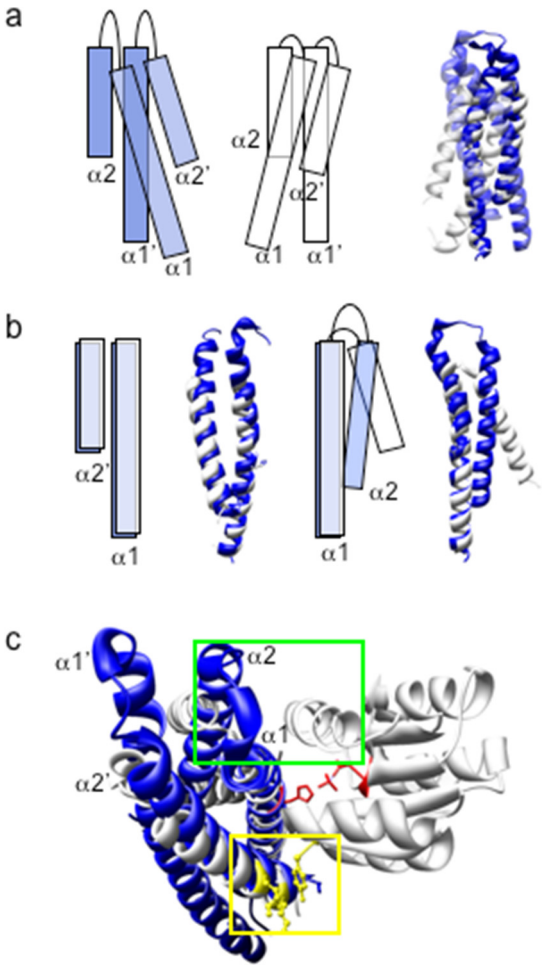


Fig. S2. Structural differences and similarities between the HisKA and Spo0B four-helix bundle domains. (*A*) Comparison between the four-helix bundle domain of SK TM0853 (blue) and Spo0B (white) reveals that the helical orientation between α1 in one monomer and α2′ in the other monomer remains unchanged, where as the orientation of α1 to α2 changes substantially (28° angle). (*B*) This becomes more evident when overlaying the different two-helices elements individually. (*C*) Overlaying the Spo0B/Spo0F cocrystal structure (white) with TM0853, it becomes apparent that the different orientation of the α2 helix in respect to α1 suggests that SK TM0853 might contact its response regulator with both helices (green box). In the Spo0B/Spo0F cocrystal structure a functionally important contact is made between the α2′ helix of Spo0B with Spo0F residue Y84 in the (β4-α4)-loop region (yellow box). The similarity in orientation of α2′ in the SK structure suggest that similar contacts should be made with its RR in the complex. The phosphotransfer residues Spo0B H30 and Spo0F D54 are depicted in red for orientation.

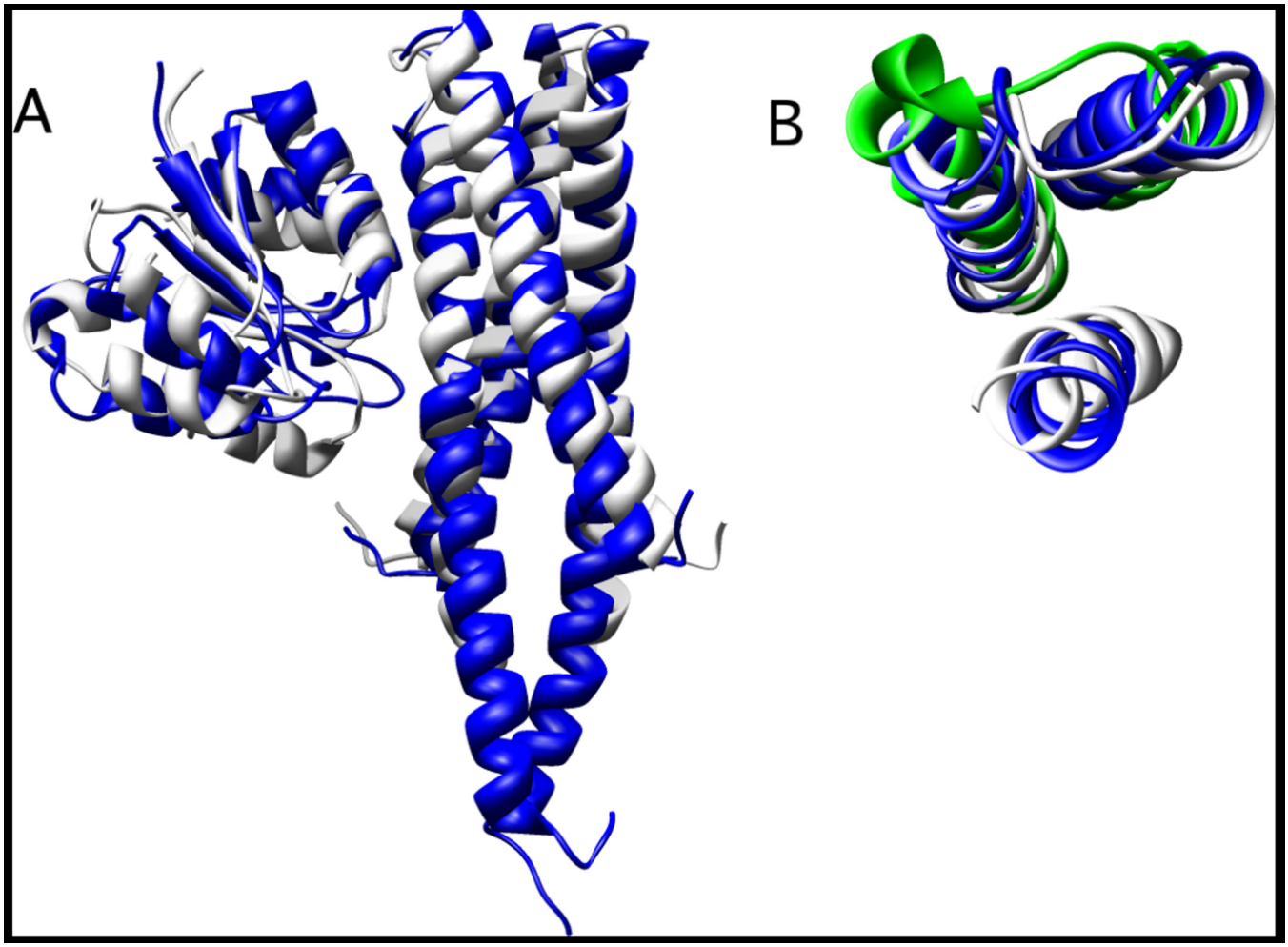
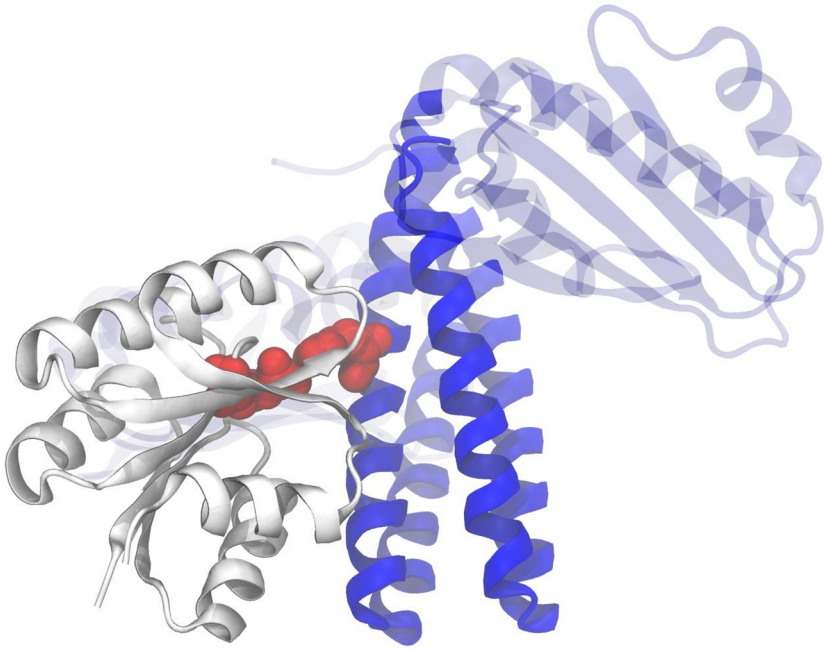


Fig. S3. Overlay of the predicted TM0853/TM0468 complex and the crystal structure. While in review the X-ray diffraction structure of the SK/RR was published. (*A*) The experimental (white) and predicted (blue) structures overlay to an RMSD of 3.3 Å. (*B*) The (1-2)-loop region of unbound SK structure (green), which was also the starting point for our simulations undergoes some conformational changes upon RR binding (white), that were almost perfectly captured in the predicted structural model (blue), demonstrating the power of our approach.



Movie S1. An exemplary structure-based simulation docks Spo0B (blue, mobile C-terminal region transparent) and Spo0F (white). His-30 and Asp-54, the phosphoryl transferring groups, are highlighted in red. The DCA predicts contacts between 6 amino acid pairs (yellow). When either the His-Asp contact or the DCA contacts are formed during the simulation, they are highlighted as solid sticks.

Movie S1 (WMV)

Other Supporting Information Files

Dataset S1